

Diseño de un amplificador operacional de transconductancia aplicando técnicas de optimización multiobjetivo

Design of an operational transconductance amplifier applying multiobjective optimization techniques

Roberto Pereira-Arroyo¹
Roberto Molina-Robles²
Alfonso Chacón-Rodríguez³

Fecha de recepción: 19 de abril del 2013
Fecha de aprobación: 28 de junio del 2013

Pereira-Arroyo, R; Molina-Robles, R; Chacón-Rodríguez, A. Diseño de un amplificador operacional de transconductancia aplicando técnicas de optimización multiobjetivo. *Tecnología en Marcha*. Vol. 27, N° 1. Pág 3-12

- 1 Escuela de Ingeniería en Electrónica, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica. Correo electrónico: rpereira@itcr.ac.cr.
- 2 Escuela de Ingeniería en Electrónica, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica. Correo electrónico: alchacon@itcr.ac.cr.
- 3 Escuela de Ingeniería en Electrónica, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica. Correo electrónico: roberc_25@hotmail.com.

Palabras clave

Optimización multiobjetivo; amplificador operacional de transconductancia; algoritmos genéticos; CMOS.

Resumen

En este trabajo, el problema en cuestión consiste en el dimensionamiento de un Amplificador Operacional de Transconductancia (OTA). El frente de Pareto se introduce como un concepto útil de análisis con el fin de explorar el espacio de diseño de este tipo de circuitos analógicos. Se emplea un algoritmo genético (GA) para detectar automáticamente este frente, en un proceso que de manera eficiente encuentra parametrizaciones óptimas y sus valores correspondientes en un espacio agregado de aptitudes. Ya que el problema es tratado como una tarea de optimización multiobjetivo, las diferentes medidas del amplificador como la transconductancia, razón de cambio de salida, el rango lineal y la capacitancia de entrada se utilizan como funciones de aptitud. Por último, se presentan los resultados de simulación, utilizando una tecnología CMOS estándar de $0,5\mu\text{m}$.

Keywords

Multiobjective optimization; operational transconductance amplifier; genetic algorithms; CMOS.

Abstract

In this paper, the problem at hand consists in the sizing of an Operational Transconductance Amplifier (OTA). The Pareto front is introduced as a useful analysis concept in order to explore the design space of such analog circuit. A genetic algorithm (GA) is employed to automatically detect this front in a process that efficiently finds optimal parameterizations and their corresponding values in an aggregate fitness space. Since the problem is treated as a multi-objective optimization task, different measures of the amplifier like the transconductance, the slew rate, the linear range and the input capacitance are used as fitness functions. Finally, simulation results are presented, using a standard $0,5\mu\text{m}$ CMOS technology.

Introducción

Este documento presenta una estrategia de optimización automatizada aplicada al proceso de diseño de un Amplificador Operacional de Transconductancia (OTA), aprovechando la potencia y versatilidad de los algoritmos genéticos (GA). Algunos enfoques atan el problema de optimización de la topología específica de un circuito y de sus parámetros, lo que hace necesaria una búsqueda relativamente exhaustiva del espacio de parámetros. (MacEachern, 1999; Hassan, Anis y Elmasry, 2005). Los algoritmos genéticos, por otro lado, funcionan en un nivel de abstracción más alto, en el que no es necesaria información específica sobre el circuito que se está optimizando; el algoritmo genético solo requiere un conjunto de valores de aptitud (por ejemplo, números reales), que representa algunos de los

parámetros del circuito tales como el consumo de energía, área de silicio o su velocidad de respuesta. El optimizador propuesto utiliza el algoritmo genético llamado PESA (Algoritmo de selección por envolvente de Pareto), y se basa en un simulador de circuitos estándar (por ejemplo, Eldo, Spectre u otra versión de Spice) para hacer frente a la complejidad de los parámetros físicos del transistor MOS y la topología del circuito (Corme & Knowles, 2000). Parámetros del circuito, como tensiones de alimentación, las corrientes de polarización, anchura y longitud de transistores, son generados por el algoritmo genético y se pasan al simulador, donde se lleva a cabo el cálculo de los valores de aptitud. Por lo tanto, el diseñador puede cambiar ya sea el algoritmo de optimización o los modelos de simulación sin mucho esfuerzo.

La herramienta de optimización se aplicó en el diseño de varios OTA, con el fin de utilizarlos para diseñar una unidad de detección del ritmo cardíaco humano, incorporados comúnmente dentro de los marcapasos. Para esta unidad, se diseñaron filtros bicuadráticos y amplificadores utilizando varios OTA.

Características de los circuitos OTA

Los OTA son dispositivos que producen una salida de corriente a partir de una entrada diferencial de tensión. Su comportamiento ideal se caracteriza por una muy alta impedancia de entrada y de salida (infinitas), tal como se muestra en la figura 1. La función de transferencia del OTA se conoce como *transconductancia* y se denota como G_m .

La estructura del OTA está compuesta por dos etapas. La primera es un amplificador con entrada diferencial, la cual provee fluctuaciones de corriente como respuesta al voltaje de entrada (V^+ y V^-). La segunda etapa está compuesta por espejos de corriente que combinan estas fluctuaciones de corriente y las redirige hacia una única salida. La tecnología CMOS se utiliza ampliamente en el diseño de OTA debido a su bajo consumo de potencia, su

capacidad para desarrollar entradas diferenciales con impedancias infinitas (incluso a lazo abierto) y la posibilidad de alcanzar valores muy bajos de transconductancia.

Es importante mencionar que la transconductancia, en función de la entrada diferencial de tensión, no es lineal. Esto es ocasionado por las ecuaciones que describen a los transistores MOSFET en sus regiones de operación (Tsividis, 2002). Como la región de inversión débil está dominada por factores exponenciales y la región de inversión fuerte por factores cuadráticos, se recomienda lograr que los transistores del diseño funcionen en la zona de inversión moderada, pues ofrece el mejor compromiso entre amplitud de la zona linealizable y el consumo de potencia, acorde con Chacón (2009).

Para poder incrementar la linealidad del OTA existen mejoras que pueden ser utilizadas en el par diferencial (Krummenacher y Joehl, 1988, Nedungadi y Viswanathan, 1984). De ellas, una versión modificada del diseño de Krummenacher (1988) implementada por Chacón (2009) fue empleada en la arquitectura del OTA sujeto de ser optimizado. La estructura de entrada del OTA se presenta en la figura 2.

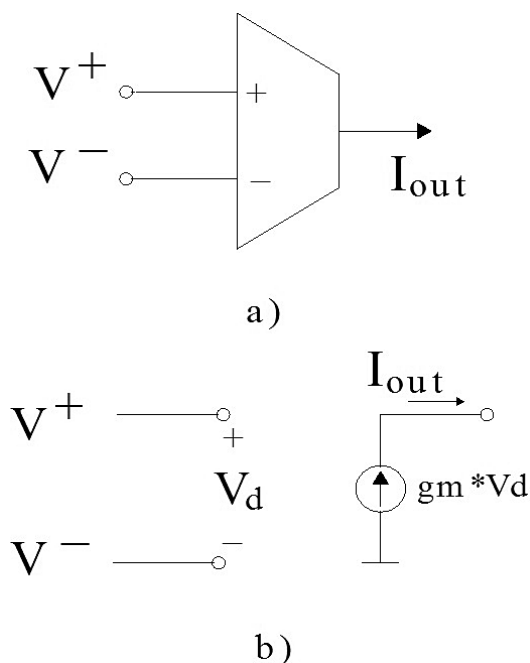


Figura 1. a) Representación de un OTA ideal. b) Circuito equivalente de pequeña señal. (Bracho, 2005)

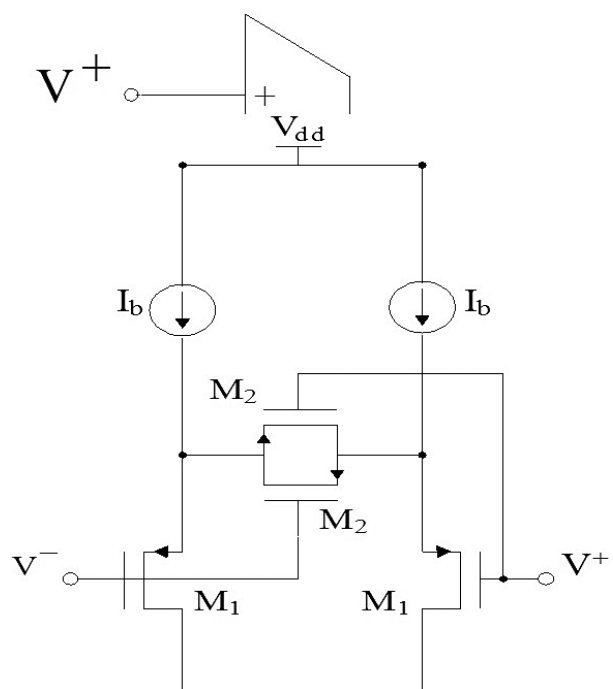


Figura 2. Etapa de entrada del OTA implementado.

La etapa de salida de corriente utiliza espejos para añadir las corrientes I_+ e I_- para obtener I_{out} . Arnaud (2004) muestra que el uso de espejos de corriente con arreglos serie-paralelo de transistores permite que la transconductancia se pueda escalar a valores más pequeños sin afectar su rango lineal y, a la vez, facilitando el trazado (*layout*) de los circuitos.

El circuito completo diseñado se muestra en la figura 3, en el que los transistores etiquetados como M_1 , M_3 y M_5 están formados por un arreglo de tres transistores unitarios conectados en serie, mientras que los transistores etiquetados M_2 son un arreglo de 18 transistores en conexión en serie también. Esto se hace con el fin de reducir los efectos de difusión lateral y mitigar los problemas de disparidad (*mismatch*) (Razavi, agosto 2000).

Los transistores M_1 se utilizan como el par de entrada diferencial; los transistores M_2 se llaman el difusor simétrico, que son responsables de la mejora de la respuesta lineal de la corriente de salida según explica Krummenacher (1988); los transistores M_3 forman los espejos de corriente que escalan la

transconductancia del circuito y los transistores M_5 se utilizan para copiar una de las ramas de corriente en el otro lado para obtener un dispositivo de terminación única. La transconductancia de salida G_m puede ser expresada, muy aproximadamente, por la siguiente ecuación:

$$G_m = \frac{g_{m1}}{m \left(1 + \frac{g_{m1}}{4g_{m2}} \right)} \quad (1)$$

donde m representa el factor de escala debido al espejo de corriente inferior; mientras g_{m1} y g_{m2} representan la transconductancia de los transistores M_1 y M_2 de la figura 3, respectivamente.

Otras ecuaciones comúnmente involucradas en diseño analógico CMOS calculado a mano se muestran en el siguiente conjunto (Tsividis, 2002, Chacón, 2009). Estas ecuaciones se derivan del modelo de EKV y se aplican a todas las regiones de operación de los transistores:

$$g_{mi} = \frac{I_{bi}}{n\phi_t} \frac{1}{f(x)} \quad (2)$$

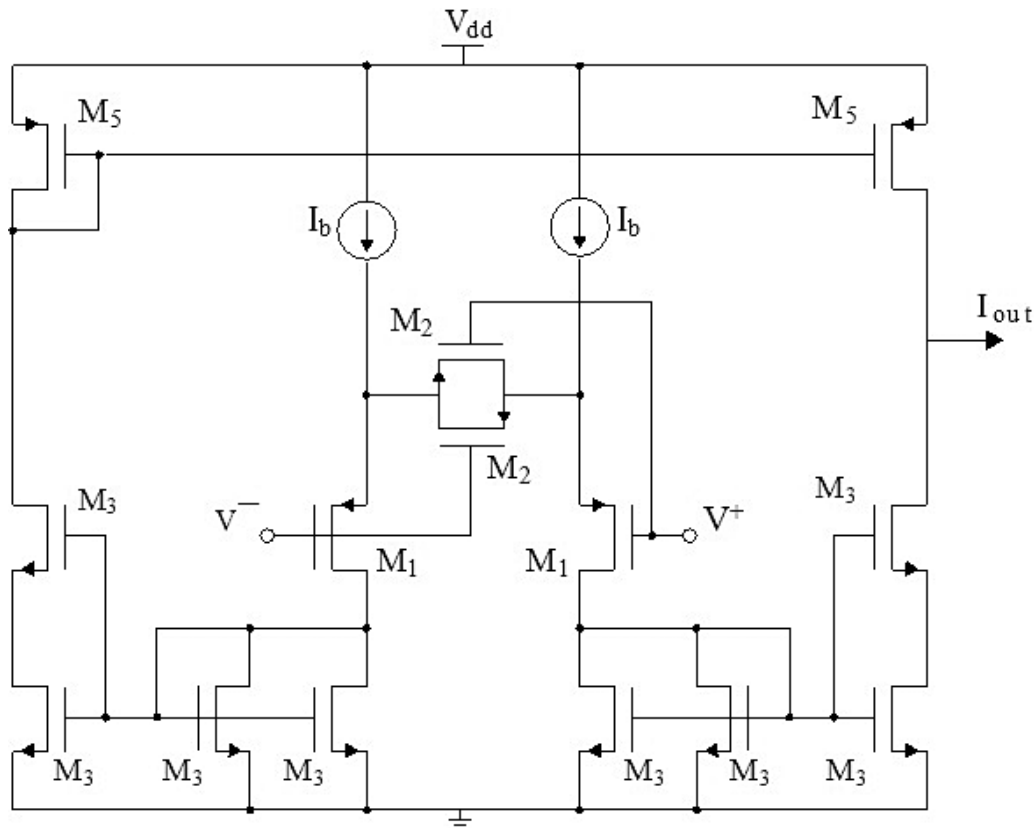


Figura 3. Esquema del circuito completo del OTA implementado.

$$f(x) = \sqrt{1 + 0.5\sqrt{x} + x} \quad (3)$$

$$x_i = \frac{I_{b_i}}{I_{z_i}} \quad (4)$$

$$I_{z_i} = 2 \frac{W_i}{L_i} C_{ox} n \phi_t^2 \quad (5)$$

donde I_{b_i} representa la corriente de polarización de CC mostrada en la figura 2, I_{z_i} es la corriente de normalización, μ , C_{ox} y n son parámetros de la tecnología de fabricación, W/L representa las dimensiones de los transistores y ϕ_t es el voltaje térmico. De las ecuaciones anteriores se puede observar que cuando el ingeniero ha de satisfacer requisitos específicos dentro de los objetivos, es necesario tener en cuenta muchos y diferentes parámetros durante el proceso de diseño. El procedimiento de diseño usual consiste en modificar algunos valores del circuito (como las dimensiones de los transistores) y, a continuación, ajustar el resto de los parámetros para cumplir los requisitos. Sin embargo, a veces este proceso tarda varias iteraciones de ensayos de prueba y error para optimizar los resultados.

La mayoría de las simulaciones y experimentos que se han ejecutado mostraron que el rango lineal ΔV es directamente dependiente de la corriente de polarización y de las dimensiones de los transistores M_1 y M_2 de la figura 3.

Filtros G_m -C: resumen y consideraciones sobre su slew rate

A. Filtros G_m -C

Los OTA se utilizan especialmente en filtros con constantes de tiempo muy grandes, que requieren alta resistencia o muy baja capacidad (Arnaud, abril 2004). El OTA presentado en la figura 3 se utilizó en el diseño de un filtro de primer orden, como el mostrado en la figura 4. El objetivo de este filtro es obtener una unidad básica y simple, que pueda ser replicada con el fin de implementar una estructura más compleja.

La función de transferencia del filtro es como se muestra en la siguiente ecuación:

$$\frac{V_{out}}{V_{in}} = \frac{1}{1 + s \frac{C}{G_m}} \quad (7)$$

y el polo donde se encuentra la frecuencia de corte se define como:

$$\omega_c = \frac{C}{G_m} \quad (8)$$

B. Razón de cambio de salida (Slew rate)

Otro parámetro importante en el diseño de filtros es su *slew rate*, ya que puede afectar al rendimiento real del ancho de banda del filtro. Se define como la máxima razón de cambio en el voltaje de salida que el filtro puede entregar. Si cualquier señal de entrada requiere una mayor tasa, entonces el filtro no será capaz de producir la respuesta de salida correcta. Esta razón (SR) para un filtro G_m -C se define como (Razavi, agosto 2000):

$$SR = \left. \frac{dV_o}{dt} \right|_{V_{o\max}} = \frac{I_{o\max}}{C} \quad (9)$$

La corriente de salida máxima posible se obtiene cuando ambos valores de I_b son dirigidos al nodo de salida. Estas corrientes pueden ser escaladas por los espejos también, así la ecuación (10) representa el cálculo de la velocidad de subida para el filtro que se muestra en la figura 4:

$$SR = \frac{2 * I_b}{m * C} \quad (10)$$

Combinando las ecuaciones (8) y (10) es posible definir la velocidad de respuesta en términos de variables de OTA y la frecuencia de corte f_c deseada.

C. Especificaciones de diseño

El diseño inicial del OTA utilizó transistores unitarios de $3\mu m/8\mu m$ para todo el circuito y una I_b de 26nA, con el fin de obtener una transconductancia de

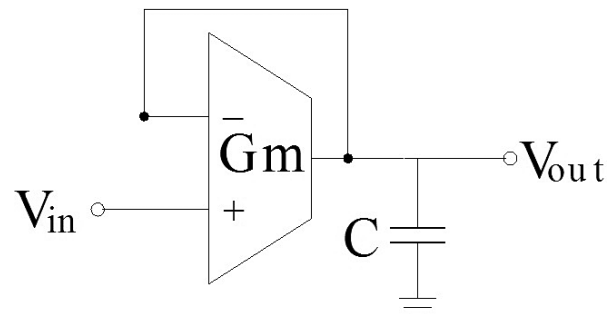


Figura 4. Esquema de un filtro G_m -C de primer orden

34nS. Con estos parámetros, los resultados obtenidos son los que se muestran en el cuadro I.

Cuadro I. Resultados de la simulación del OTA inicial.

Medida	Valor
Máximo G_m (nS)	36,57
Rango Lineal ΔV (mV)	± 260
Slew rate (mV/ μ s)	1,954
Consumo de Potencia (nW)	174,93
Capacitancia de Entrada (fF)	267,79

El objetivo de la herramienta era reducir el valor de transconductancia y la capacidad de entrada del OTA y al mismo tiempo aumentar el rango lineal al menos hasta ± 500 mV. El *slew rate* del circuito debe ser 2,75 mV/ μ s para el correcto funcionamiento del filtro en el que este OTA se aplicará. A fin de lograr esto, la herramienta de optimización ha trabajado con los siguientes parámetros del circuito: I_b , L_1 , L_2 , L_3 , L_5 , W_1 , W_2 , W_3 y W_5 . En la figura 5 se presenta el diagrama de bloques de la herramienta de optimización diseñada para resolver el problema en cuestión. Los objetivos impuestos al correr la herramienta fueron: a) aumentar el rango lineal y velocidad de respuesta, y b) disminuir tanto la capacidad de entrada como la transconductancia.

Algoritmos genéticos para la optimización multiobjetivo de circuitos

La función de aptitud agregada F para un circuito A con una parametrización u se define como:

$$F(A_u) = \Phi(f_1(A_u), f_2(A_u), \dots, f_n(A_u)) \quad (12)$$

donde cada función de aptitud individual $f_i(A_u)$ se define como monótonamente creciente con la aptitud que describe el comportamiento del circuito. En otras palabras, el algoritmo genético optimiza los parámetros de aptitud buscando el valor más alto posible para cada uno de ellos. Así que para este caso era necesario que el rango lineal y la velocidad de subida fueran tan altos como fuese posible, y por lo tanto sus valores de aptitud respectivos resultaron directamente proporcionales. Por el contrario, la capacidad de entrada y la transconductancia debían ser lo más bajos posibles, y por lo tanto sus valores de aptitud resultaron inversamente proporcionales. Las funciones f_i abarcan un espacio de aptitud multidimensional, donde cada punto representa el rendimiento de un circuito parametrizado con un punto u en el espacio de parámetros.

La forma general de Φ se supone desconocida, pero tiene que aumentar monótonamente con valores crecientes de todas las funciones de aptitud f_i . Esta condición garantiza que un punto en el espacio de aptitudes puede ser considerado más apto que todos los demás puntos cuyos valores sean menores en todas las dimensiones. En la figura 6, por ejemplo,

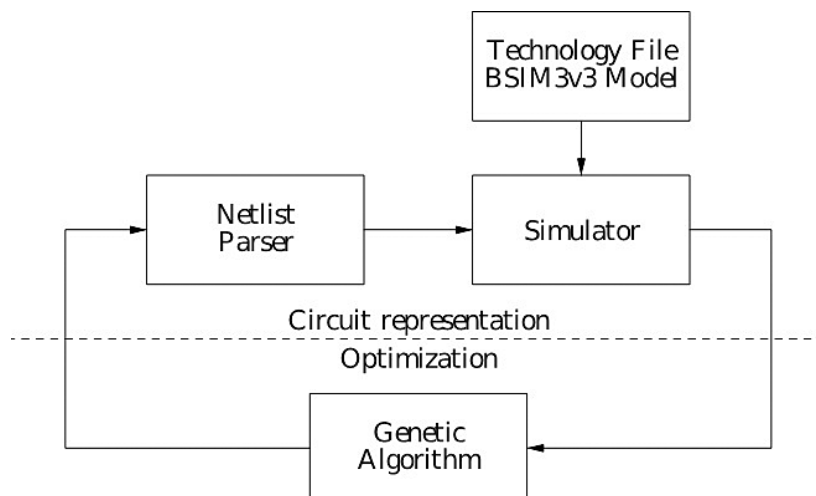


Figura 5. Herramienta de optimización basada en algoritmos genéticos.

el punto q_1 es más apto que el punto q_4 y todos los otros elementos dentro del rectángulo gris. En este contexto, el punto q_1 se dice que domina a q_4 . Todos los puntos no dominados en un conjunto definen el frente de Pareto de ese conjunto. En el ejemplo de la figura 6 este frente está definido por los puntos q_1, q_2 y q_3 . Elegir una parametrización que no está en el frente es siempre una mala elección, ya que hay otro punto en el frente con una mejor aptitud agregada. Los conceptos anteriores se pueden expresar matemáticamente mediante la ecuación siguiente:

$$\hat{P} = \left\{ u \in P_A : f(A_u) \mid \neg \exists v \in P_A : f(A_v) \succ f(A_u) \right\} \quad (13)$$

donde \hat{P} es el frente Pareto, f es el vector de funciones de aptitud $[f_1, \dots, f_n]^T$ y P_A es el espacio de parámetros del circuito A. La relación de orden parcial ' \succ ' en f describe la propiedad de dominación definida como:

$$f(A_v) \succ f(A_u) \Leftrightarrow \forall_i : f_i(A_v) \geq f_i(A_u) \wedge \exists_j : f_j(A_v) > f_j(A_u) \quad (14)$$

El proceso de evaluación puede por lo tanto ser considerado como un proceso de mapeo que transforma el espacio de parámetros válido P_A en una región conectada en el espacio de *fitness* $[f_1, \dots, f_n]^T$. El frente de Pareto es la frontera de esta región delimitada por los óptimos parciales (Müller-Gritschneider, junio 2009). Cualquier algoritmo que encuentra el frente de Pareto para un conjunto de puntos de forma física implementa las ecuaciones

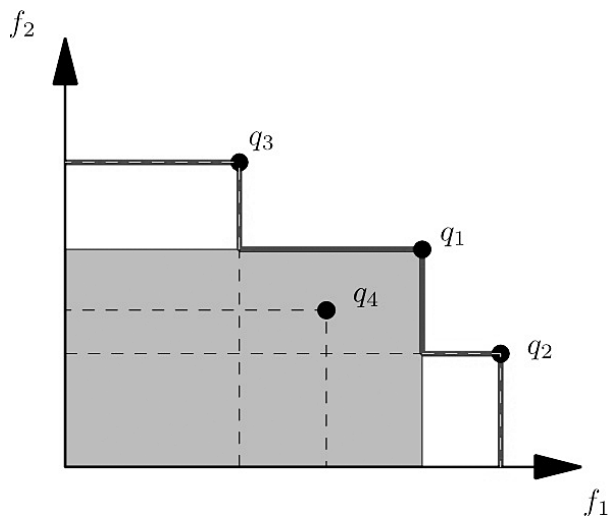


Figura 6. Frente de Pareto.

(13) y (14). Dado que el espacio de parámetros P_A generalmente contiene un número infinito de parametrizaciones, el siguiente problema consiste en la elección de un conjunto representativo de muestras de P_A , de tal manera que su frente de Pareto puede suponerse que es una aproximación fiable del frente correspondiente al espacio de diseño completo. Un enfoque ingenuo sería muestrear periódicamente los valores de cada parámetro, dado que el número de evaluaciones necesarias aumentará exponencialmente con el número de parámetros. Por ejemplo, un circuito con siete parámetros (variables de diseño), cada uno muestreado cinco veces, requeriría $5^7 = 78,125$ evaluaciones. Para evitar esta búsqueda por fuerza bruta de parámetros, se emplea el algoritmo evolutivo multiobjetivo PESA. Este enfoque genético suprime el cálculo de parametrizaciones inútiles y se concentra en el análisis de aquellas regiones del espacio de parámetros que proporcionan resultados prometedores. Incluso si este algoritmo también discretiza el espacio de parámetros, a través de una representación numérica con un número finito de bits, la resolución utilizada para cada parámetro puede ser tan alta como sea necesaria, sin la amenaza de una explosión exponencial del espacio de búsqueda. El número de evaluaciones requeridas es entonces proporcional al número de bits utilizados para representar una parametrización.

Resultados

La figura 7 muestra el frente de Pareto tridimensional del OTA previamente mostrado. Este frente contiene 1500 elementos (parametrizaciones) y fue generado por el algoritmo genético PESA. El gráfico muestra el compromiso entre las tres funciones de aptitud, donde el tratar de maximizar una de ellas, implica disminuir las otras.

El cuadro 2 contiene una lista de algunos de los resultados seleccionados dados por la herramienta de optimización. Estos resultados contienen casos con amplio rango lineal, transconductancias bajas y capacitancias bajas. La condición de óptimo *slew rate* se logra con bajos valores de G_m .

La figura 8 presenta los resultados de la simulación para el mejor de los casos que se encontró en esta corrida. La figura muestra la respuesta de corriente de salida para un barrido de tensión de entrada de $-1V$ a $+1V$.

El cuadro 3 muestra una comparación de las simulaciones obtenidas para el caso del mejor OTA y el OTA inicial. Cuando se reduce la transconductancia, por ende se reduce el tamaño de los capacitores (para mantener el polo en el mismo lugar), produciendo una reducción del área del circuito. Todas las otras especificaciones de diseño también fueron

mejoradas: se redujeron la capacidad de entrada y el consumo de energía, mientras que el intervalo lineal y la velocidad de respuesta se han mejorado en gran medida. Finalmente, el cuadro 4 contiene las dimensiones de los transistores unitarios propuestos como resultado de correr la herramienta de optimización.

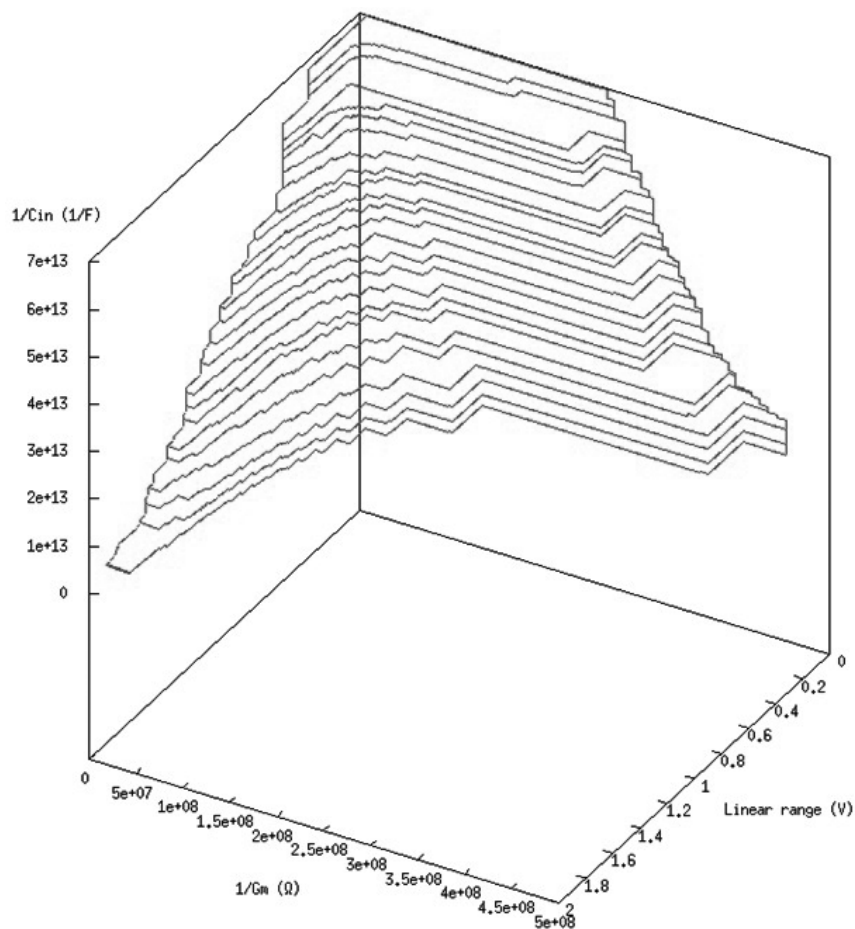


Figura 7. Frente de Pareto del OTA diseñado. El gráfico contiene tres aptitudes: capacidad de entrada, transconductancia y rango de tensión lineal.

Cuadro 2. Datos representativos dados por la herramienta de optimización.

G_m (nS)	ΔV (mV)	C_{in} (fF)	Slew rate (mV/ μ s)	I_b (nA)
33,55	± 949	235,58	2,131	87,226
40,90	± 581	86,35	1,748	64,871
44,80	± 624	93,40	2,054	52,097
28,90	± 549	94,63	2,473	39,323
18,0	± 504	85,74	3,971	26,548
16,55	± 506	88,64	4,319	20,161

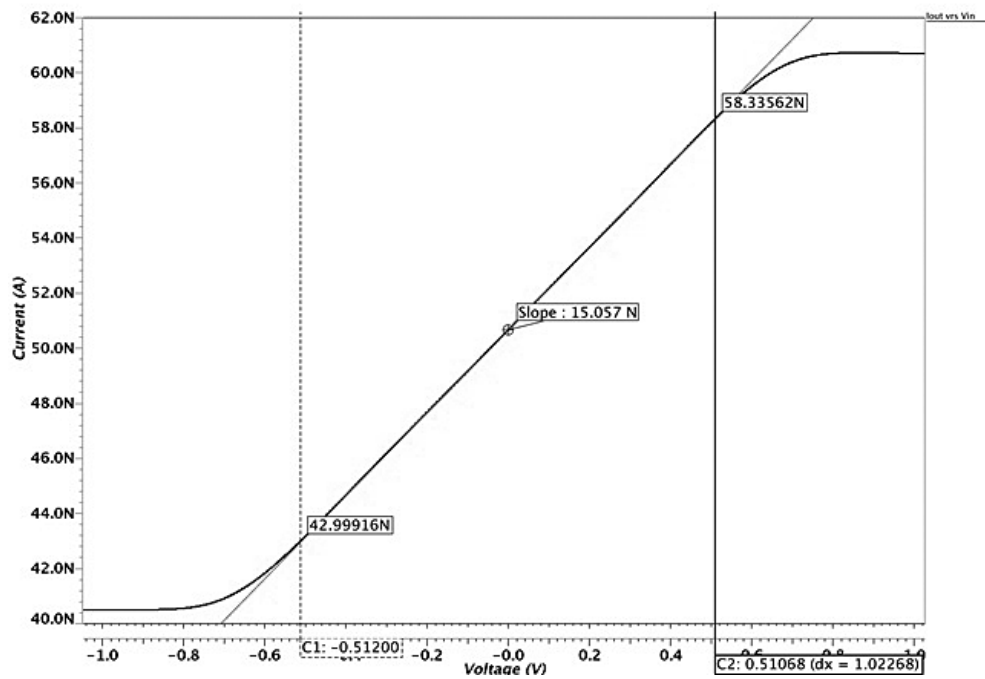


Figura 8. Curva de corriente de salida como una función de la tensión de entrada. La pendiente de esta curva da la transconductancia del circuito.

Cuadro 3. Resultados de la simulación del mejor OTA obtenido.

Medida	Mejor OTA	OTA Inicial
Máximo G_m (nS)	15,077	36,57
Rango Lineal ΔV (mV)	± 512	± 260
Slew rate (mV/ μ s)	3,676	1,954
Consumo de Potencia (nW)	144,3	174,93
Capacitancia de Entrada (fF)	89,63	267,79

Cuadro 4. Dimensiones de transistores unitarios del mejor OTA encontrado.

Parámetro	Valor
L_1 (μ m)	7,2
L_2 (μ m)	8
L_3 (μ m)	2,1
L_5 (μ m)	2,1
W_1 (μ m)	1,5
W_2 (μ m)	1,5
W_3 (μ m)	5,1
W_5 (μ m)	5,1

Conclusiones

Este trabajo presenta la aplicación de una estrategia automatizada con optimización multiobjetivo para diseñar y mejorar un circuito de OTA. La herramienta EDA (Electronic Design Automation) está destinada a ayudar en la puesta a punto del circuito y el proceso de mejora a través de la generación de parametrizaciones optimizadas.

Este mecanismo no solo reduce en gran medida el tiempo necesario para diseñar y simular este tipo de circuito sino que también permite una solución óptima. El diseño basado en cálculo a mano se queda atrás frente a estas ventajas, aun con menor precisión que las herramientas automatizadas. Sin embargo, los criterios de selección del mejor de los casos sigue siendo trabajo del diseñador, ya que el frente de Pareto ofrece un mapa de puntos optimizados.

Una comparación entre los valores del cuadro 3 muestra cómo la herramienta EDA efectivamente ha mejorado el funcionamiento del circuito en cada aspecto que se pretende conseguir.

Bibliografía

- Arnaud, A. (abril 2004). *Very large time constant gm-c filters*. (Tesis Doctoral). Instituto de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay.
- Baru, M. (2009). Analog design methodologies when every nanoampere counts. En: *CMOS ET Workshop Vancouver*.
- Bracho, S. (2005). Amp.operacionales de transconductancia. Recuperado de: <http://www.teisa.unican.es/teisa/classes/docencia/bracho/I685/AmpOp.pdf>
- Chacón, A. (2009). *Circuitos integrados de bajo consumo para detección y localización de disparos de armas de fuego*. (Tesis doctoral). Facultad de Ingeniería, Departamento de Ingeniería Electrónica, Universidad Mar del Plata, Mar del Plata, Argentina.
- Corne, D. & Knowles, J. (2000). The pareto envelope-based selection algorithm for multiobjective optimization. En: *PPSN VI: Proceedings of the International Conference on Parallel Problem Solving from Nature*, 839-848.
- Hassan, H., Anis, M. & Elmasry, M. (2005). MOS current mode circuits: analysis, design, and variability. *IEEE Trans. VLSI Syst.* 13(8): 885-898.
- Krummenacher, F. & Joehl, N. (1988). A 4-mhz cmos continuous-time filter with on-chip automatic tuning. *IEEE J. Solid-State Circuits* 23(3): 750-758.
- MacEachern, L.A. (1999). Constrained circuit optimization via library table genetic algorithms. En: *Proc. IEEE Int. Symp. Circuits and Systems ISCAS '99*, 6, 310-313.
- Müller-Gritschneider, D. (junio 2009). Deterministic performance space exploration of analog integrated circuits considering process variations and operating conditions. (Tesis doctoral). Technical University Munich, Munich, Alemania.
- Nedungadi, A. & Viswanathan, T. (1984). Design of linear cmos transconductance elements. *IEEE Trans. Circuits Syst.* 31(10): 891-894.
- Razavi, B. (agosto 2000). *Design of Analog CMOS Integrated Circuits*. 1a. ed. McGraw-Hill Science/Engineering/Math.
- Tsividis, Y., Czarml, Z. & Fang, S.C. (1986). MOS transconductors and integrators with high linearity. *Electronics Letters* 22(5): 245-246.
- Tsividis, Y. (2002). *Mixed Analog-Digital VLSI Devices and Technology*. Columbia University, USA: World Scientific Publishing.